INCORPORATING YEARLY DERIVED WINTER WHEAT MAPS INTO WINTER WHEAT YIELD FORECASTING MODEL

S. Skakun[1,2], *B. Franch*[1,2], *J.-C. Roger*[1,2], *E. Vermote*[2],
I. Becker-Reshef[1], *C. Justice*[1], *A. Santamaría-Artigas*[1]

[1]Department of Geographical Sciences, University of Maryland, College Park MD 20742, USA
[2]NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA

ABSTRACT

Wheat is one of the most important cereal crops in the world. Timely and accurate forecast of wheat yield and production at global scale is vital in implementing food security policy. Becker-Reshef et al. (2010) developed a generalized empirical model for forecasting winter wheat production using remote sensing data and official statistics. This model was implemented using static wheat maps. In this paper, we analyze the impact of incorporating yearly wheat masks into the forecasting model. We propose a new approach of producing in season winter wheat maps exploiting satellite data and official statistics on crop area only. Validation on independent data showed that the proposed approach reached 6% to 23% of omission error and 10% to 16% of commission error when mapping winter wheat 2-3 months before harvest. In general, we found a limited impact of using yearly winter wheat masks over a static mask for the study regions.

Index Terms— Winter wheat, prediction, crop mapping, MODIS

1. INTRODUCTION

Wheat is one of the most important crops in the world. Timely and accurate prediction of wheat yield and wheat production at global scale still remains a challenge from both scientific and operational perspectives [1]. Remote sensing data from space enable monitoring and quantification of crop state at global scale. Biophysical parameters and vegetation indices derived from Earth observation (EO) satellites are used to assimilate data into crop growth models [2]–[4] or used as proxies in empirical models to predict crop yield [5]–[9]. Many models have been proposed so far at regional and global scale. Becker-Reshef et al. (2010) [5] developed a generalized empirical model for forecasting winter wheat yield and production using remote sensing data and official statistics. The main pillar of the model is utilization of a relationship between the yield, the seasonal peak normalized difference vegetation index (NDVI) derived from Moderate resolution Imaging Spectro-radiometer (MODIS) and maximum winter wheat percentage (*Mpct*) per corresponding administrative units of the country. The model was further improved by Franch et al. (2015) [8] to incorporate Growing Degree Day (GDD) information into the model to improve the timeliness of the forecasts. The model was applied for multiple countries including USA, Ukraine and China, and showed ability to perform accurate forecasts of wheat production at national/state level 2-2.5 months before harvest. However, it should be noted that these models were using static winter wheat masks for corresponding countries. At present, winter wheat maps are not available at global scale and produced regularly for only selected countries [10]. Also, in many cases these maps are produced after the growing season making it difficult to use in crop forecasting models.

In this paper, we explore the value of adding winter wheat masks derived on yearly basis in winter wheat yield forecasting models [5], [8]. An automatic method for mapping winter crops based on satellite imagery and official statistics is proposed. A theoretical error budget for the forecasting model is calculated in order to estimate uncertainties associated with the model's components.

2. MATERIALS

The study was performed for the state of Kansas, as one of the main winter wheat producing states in the U.S., and Ukraine. The winter wheat yield and production forecasting model was built using the following set of datasets: BRDF-corrected MODIS surface reflectance time series data [11] at daily temporal resolution and 0.05° spatial resolution (5600 m at the equator) of Climate Modeling Grid (CMG); official statistics at administrative units (Agricultural Statistics Districts for Kansas and oblast level for Ukraine); and winter wheat masks.

For winter wheat mapping, we used NDVI temporal profiles extracted from the MODIS MOD13Q1 product. MOD13Q1 images are provided every 16 days at the 250 m spatial resolution in the sinusoidal projection. NDVI images are composited over a 16-days interval to create a cloud-free map with minimal atmospheric and sun-surface-sensor angular effects [12]. Data for the territory of Kansas and

Ukraine and 2000–2014 time interval were downloaded from the Land Processes Distributed Active Archive Center (LPDAAC) of the USGS (https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table). The MOD13Q1 NDVI composites were further converted to the Albers Equal Area projection.

We also used Cropland Data Layer (CDL) data for Kansas for 2006–2014 in order to validate the proposed approach on in season winter wheat crop mapping.

3. METHODOLOGY
3.1. Winter wheat yield and production forecasting

The Becker-Reshef et al. (2010) method [5] is based on the assumption that the yield is positively and linearly correlated to the seasonal maximum NDVI (adjusted for background noise) at the administrative unit (AU, county or oblast) level and to the purity of the wheat signal. Thus, they developed a regression model that was calibrated and applied at the state level in Kansas and was proven directly applicable at the national level in Ukraine. Seasonal maximum NDVI for AU is computed in the following way. First, daily average NDVI is calculated over the 5% purest winter wheat pixels at 0.05°. Then, the derived NDVI is adjusted for bare soil by subtracting the minimum 5% values for the years studied and seasonal maximum adjusted NDVI is computed (MA_NDVI). The model utilizes a generalized relationship $S = 9.61 + (0.05 Mpct)$, where $Mcpt$ is the weighted average of the percent wheat values of the purest 5% wheat dominated pixels for each AU, to compute the slope and the derived MA_NDVI value to predict the yield:

$Yield = S * MA\_NDVI$.

This method was extended by Franch et al. (2015) [8] to incorporate GDD information to improve timeliness of the model. In particular, current and historical GDD and NDVI values are used to predict the NDVI peak, thus enabling to obtain forecasts of the yield and production 1-1.5 months before NDVI peak.

A theoretical error budget for the model was analyzed to estimate uncertainties associated with the model's components. The main components of the model are as follows: NDVI derived from MODIS, winter wheat percentage and coefficients of the relationship between the S and $Mpct$. Corresponding partial derivatives for each model component were calculated and the final error of the yield was derived. Therefore, such approach allowed us to analyze contribution of each component into the final yield uncertainties. We found that the derived theoretical error of winter wheat yield forecasting at 0.05° resolution was 8.5%–18.3% that was consistent with previously derived results when comparing to official statistics [5].

3.2. In season winter crops mapping
3.2.1. Method description

The winter wheat forecasting model was developed and implemented using static winter wheat maps. In this study, we analyze whether the use of dynamic yearly winter wheat maps would improve the model while preserving its main advantageous characteristics: little data input requirements, applicability at global scale and timeliness.

Yearly winter wheat maps are not available at global scale, and produced regularly only for select countries (e.g. CDL for USA). Therefore, it is necessary to provide an approach that can be applied at global scale with little requirements to input data, and can produce in season winter wheat masks (2-2.5 months before harvest) in order to respect timeliness capabilities of the forecasting model.

The following approach is proposed in this paper that is based on NDVI derived from MOD13 product at 250 m spatial resolution, official statistics on planted crop area at AU and cropland mask. The approach is based on the phenological development of winter wheat that has the developed biomass in late autumn and early spring while other crops have no biomass within these time periods. For each pixel at 250 m resolution, we derive two features (hereafter denoted as "Fall" and "Spring") from MODIS NDVI time-series that exhibit maximum NDVI values during two time periods: autumn of the previous year and early spring in the current year. The derived maximum NDVI values are also subtracted from 15 years average minimum NDVI during the crop growth season to account for soil background.

In order to identify winter wheat pixels corresponding thresholds for derived features are to be calculated. Only cropland pixels were taken into account. Thresholds for features to discriminate between winter crops and non-winter crops were derived by calculating the area for corresponding AUs and matching it to official statistics on planted areas. These thresholds were calculated separately for each AU in order to account for regional peculiarities of wheat growth.

3.2.2. Method implementation and validation

This approach was applied for Kansas and Ukraine to derive yearly winter wheat masks for 2001-2014. Maximum NDVI values for deriving the "*Fall*" feature were selected from 31-Oct to 31-Dec for Kansas and Ukraine, and maximum NDVI values for extracting the "*Spring*" feature were selected from 22-Mar to 25-May for Ukraine, and from 06-Mar to 22-Apr for Kansas. These time periods were selected taking into account several factors, in particular availability of MODIS images and phenological development of winter wheat comparing to other crop types in the study areas. Thresholds for "*Fall*" and "*Spring*" features were calculated in such a way to match official statistics on planted winter wheat areas at ASD level in Kansas and oblast level in Ukraine.

Usually official statistics on crop area serves as a source for validating the derived maps. However, this cannot be applied in our case since official statistics was used for calculating and calibrating thresholds for identifying winter wheat pixels. Therefore, CDL maps for Kansas were used to independently validate the proposed approach.

4. RESULTS

The winter wheat masks for Kansas derived using the proposed approach were compared to the CDL derived masks at different scales: (i) at CDL spatial resolution to estimate omission error (OE) and commission error (CE) of the derived maps (Fig. 1); (ii) at 0.05° resolution to estimate the differences of wheat proportions at this scale from the proposed approach and CDL (Fig. 2).

When comparing the winter wheat masks produced with the proposed approach to the CDL-derived masks, the average over the 2006-2014 period omission error (OE) was 13% and commission error (CE) was 11%. The best values were obtained for 2006 with OE of 6.1% and CE of 9.8%, and the worst values were obtained for 2013 with OE of 23.4% and CE of 15.4%. Fig. 1 shows comparison of winter wheat masks derived from CDL and using the proposed approach depending on the winter wheat purity pixel at 250 m resolution. Blue line in Fig. 1 shows theoretical Pareto boundary derived from 56 m (2006) CDL maps. The CDL maps for winter crops were converted to the 250 m proportional maps. For each threshold that defines winter crop, a low resolution 250 m binary map was derived and compared to high resolution map (at 56 m resolution), and a pair OE/CE was estimated. Red line in Fig. 1 shows the Pareto boundary for the derived maps. Different points were derived utilizing different thresholds for determining winter crops purity at 250 m resolution.
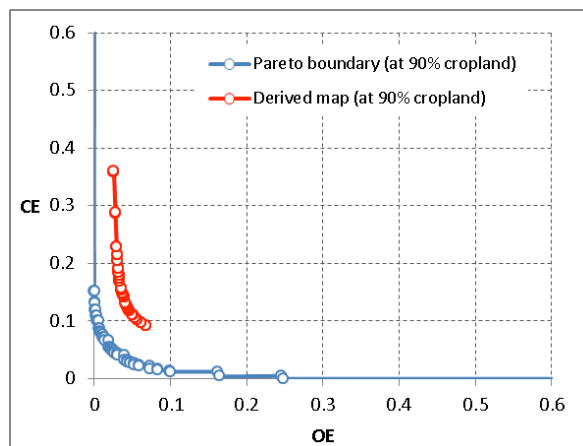

Fig. 1. Comparison of OE and CE values for the CDL mask (blue line) and derived mask (red line) for Kansas in 2006 depending on the winter wheat purity pixel at 250 m resolution.

The main causes of errors were spatial discrepancy between the derived maps at 250 m resolution and the reference CDL at 56 m or 30 m (after 2010) resolution, and

extreme meteorological conditions that forced considerable delays in wheat development and therefore inability to detect winter crops from satellite imagery during the used autumn or spring periods. In particular, in 2013 Kansas experienced lack of precipitation and freezing temperatures (especially in western part) that caused only 27% of crops to be in good to excellent by the end of April where the "Spring" features were selected.

Comparison of winter wheat proportions at 0.05° resolution derived from CDL and our maps showed small difference: average over the 2006-2014 period root mean square error (RMSE) was 4.75% and R squared was 0.9.
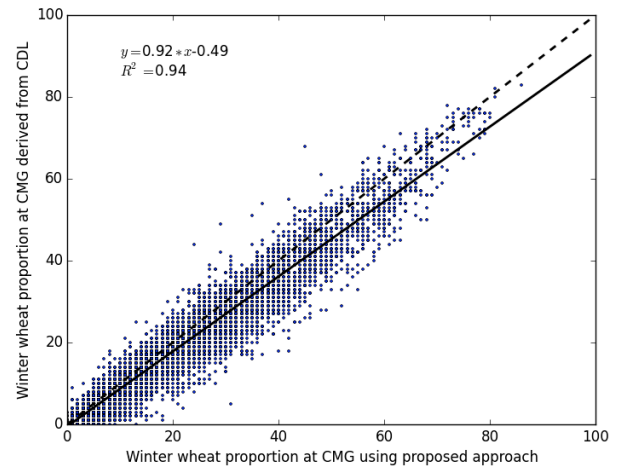

Fig. 2. Comparison of winter wheat proportions for Kansas for 2006 at 0.05° resolution using CDL and the proposed approach.
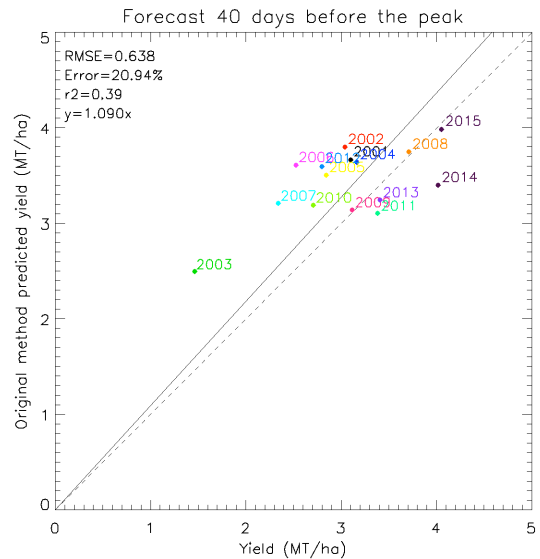

Fig. 3. Winter wheat yield predictions versus official statistics for Ukraine for 2001-2015 using a GDD and yearly masks incorporated model.

The derived yearly winter wheat maps for Ukraine were further incorporated into the forecasting model to predict the winter wheat yield for Ukraine at national level. Fig. 3

shows the comparison of the derived results for 2001-2015 versus official statistics when applying the model with GDD information, i.e. 40 days before NDVI peak. The obtained RMSE between satellite-derived forecasts and official statistics at national level was 0.64 t/ha. This accuracy was achieved around May 5th, i.e. 40 days before NDVI peak for winter wheat, which is 2-2.5 months before harvest that takes place, as a rule, in late June or early July.

In general, we found a limited impact of using yearly maps over a static map suggesting little contribution at 0.05° resolution to the final yield error. This was also observed in Kansas when applying the CDL derived winter wheat masks. This approach will be further applied to other major wheat regions to identify on the impact of using yearly wheat masks in the forecasting model.

5. CONCLUSIONS

In this paper, we explored the impact of using yearly winter wheat masks when averaging NDVI in the winter wheat yield forecasting model operating at 0.05° resolution. For this, we proposed a new approach for in season winter wheat mapping that is relied on satellite imagery and official statistics data only. The approach is based on manually constructed features to capture dynamics of winter wheat comparing to other crop types and official statistics on administrative units to calibrate feature values. The developed approach has the following advantages: little input data requirements; potential to be applicable at global scale in automatic way; and in season capabilities. Disadvantages include: failing to identify winter wheat fields in case of considerable delays in crop growth. The derived yearly masks were incorporated into the winter wheat yield forecasting model for Ukraine and Kansas. We found a limited impact of using yearly winter wheat masks over a static mask for the study regions.

6. REFERENCES

[1] I. Becker-Reshef, C. Justice, M. Sullivan, E. Vermote, C. Tucker, A. Anyamba, J. Small, E. Pak, E. Masuoka, J. Schmaltz, M. Hansen, K. Pittman, C. Birkett, D. Williams, C. Reynolds, and B. Doorn, "Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project," *Remote Sensing*, vol. 2, pp. 1589–1609, 2010.

[2] Y. Curnela, A.J.W. de Wit, G. Duveiller, and P. Defourny, "Potential performances of remotely sensed LAI assimilation in WOFOST model based on an OSS Experiment," *Agricultural and Forest Meteorology*, vol. 151, no. 12, pp. 1843–1855, 2011.

[3] A.J.W. de Wit, G. Duveiller, and P. Defourny, "Estimating regional winter wheat yield with WOFOST through the assimilation of green area index retrieved from MODIS observations," *Agricultural and Forest Meteorology*, vol. 164, pp. 39–52, 2012.

[4] H. Fang, S. Liang, and G. Hoogenboom, "Integration of MODIS LAI and vegetation index products with the CSM-CERES-Maize model for corn yield estimation," *International Journal of Remote Sensing*, vol. 32, no. 4, pp. 1039–1065, 2011.

[5] I. Becker-Reshef, E. Vermote, M. Lindeman, and C. Justice, "A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data," *Remote Sensing of Environment*, vol. 114, no. 6, pp. 1312–1323, 2010.

[6] F. Kogan, L. Salazar, and L. Roytman, "Forecasting crop production using satellite-based vegetation health indices in Kansas, USA," *International Journal of Remote Sensing*, vol. 33, no. 9, pp. 2798–2814, 2012.

[7] F. Kogan, N. Kussul, T. Adamenko, S. Skakun, O. Kravchenko, O. Kryvobok, A. Shelestov, A. Kolotii, O. Kussul, and A. Lavrenyuk, "Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models," *International Journal of Applied Earth Observation and Geoinformation*, vol. 23, pp. 192-203, 2013.

[8] B. Franch, E.F. Vermote, I. Becker-Reshef, M. Claverie, J. Huang, J. Zhang, C. Justice, and J.A. Sobrino, "Improving the timeliness of winter wheat production forecast in the United States of America, Ukraine and China using MODIS data and NCAR Growing Degree Day information," *Remote Sensing of Environment*, vol. 161, pp. 131–148, 2015.

[9] R. López-Lozano, G. Duveiller, L. Seguini, M. Meroni, S. García-Condado, J. Hooker, O. Leo, and B. Baruth, "Towards regional grain yield forecasting with 1 km-resolution EO biophysical products: Strengths and limitations at pan-European level," *Agricultural and Forest Meteorology*, vol. 206, pp. 12–32, 2015.

[10] USDA National Agricultural Statistics Service Cropland Data Layer. 2015. Published crop-specific data layer [Online]. Available at https://nassgeodata.gmu.edu/CropScape/ (accessed 18/12/2015). USDA-NASS, Washington, DC.

[11] E. Vermote, C.O. Justice, and F.M. Breon, "Towards a generalized approach for correction of the BRDF effect in MODIS directional reflectances," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 3, pp. 898–908, 2009.

[12] A. Huete, C. Justice, and W. van Leeuwen. 1999. MODIS Vegetation Index (MOD13) Algorithm Theoretical Basis Document (ATBD). Version 3 [Online]. Available at http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf (accessed 18/12/2015).